Assigned Proposal Number: 73

Strategic University Research Partnership
Director's Research and Development Fund Proposal FY 2009
Due Date: December 12, 2008, by 4 PM PST

1. Title
Strategic technology research to improve the utilization of assimilated ocean state estimates in future mission design and planning and to seed new collaborations in this area.

2. JPL Principal Investigator
Dimitris Menemenlis (3248)
menemenlis@jpl.nasa.gov

3. Co-Investigator(s)
Christopher Hill
Massachusetts Institute of Technology
Earth, Atmospheric, and Planetary Sciences
cnh@mit.edu

4. Total Budget Request for FY09:
$99,917.00

5. Student Participants
Constantinos Evangelinos
ce107@MIT.EDU
Post Doc

6. Identify one Primary (P) Topic Area to which your proposal applies and any Secondary (S) Area(s).

P	S	Topical Area
		1. Advance Solar System Exploration in New Directions: To Understand Planetary Formation and Evolutionary Pathways, and to Seek, Discover and Inventory the Organic Materials in the Solar System and Elucidate Their Origins
		2. Determining the Geometry and Structure of Our Universe
		3. Characterizing Exoplanets Where Life Could Exist
X		4. An Integrated Earth System Science Information System for Research and Applications
		5. Achieving Breakthrough Increases in Interplanetary Communications
		6. Enabling Robotic Missions to Scientifically Interesting Extreme Environments
		7. Enabling Autonomous Human Missions to the Moon and Mars
		8. Fractionated, Distributed, Repairable, Reconfigurable, Reusable Missions
		9. Large, Precise Space Structures to Enable Future Observing Instruments
		10. Develop and Nurture the Next Generation of "Rocket Scientists"

7a. General Objectives
We propose to research technology that will directly enhance the use of high-resolution assimilated ocean state estimates and associated modeling assets in future mission design. The context of this work is developing and deploying prototype technology that will allow scientists and engineers to more readily and routinely incorporate high spatial and temporal resolution ocean state information into mission design in addition to enhancing existing science research activities.

There are currently several collaborations (notably the ECCO-GODAE and ECCO2 projects) between MIT and JPL in the area of ocean state estimation. The basic methodology underlying these projects centers on using NASA satellite data, together with in-situ and remotely sensed data from other sources, to formally constrain and optimize ocean simulation trajectories by adjusting uncertain parameters, e.g., boundary conditions and initial conditions for these simulations. The outcomes of these efforts are state estimates for the full depth, time-dependent ocean state over periods of one to a few decades with unprecedented temporal and spatial resolution. To date, these collaborations have focused on developing science insight into the working of the planetary ocean and climate system. However, the high resolution state estimates now beginning to emerge from the ECCO2 project (see http://ecco2.org) are also, potentially, well suited to contributing to the design and evaluation of future ocean and cryosphere related missions, including for example, the Surface Water and Ocean Topography (SWOT) mission, the AQUARIUS mission, and the DESdynI mission. The work proposed here aims to prototype a key technological element of developing new capabilities, as a step toward making mission team interactions a more routine aspect of the ECCO2, and of future monitoring and measuring program, state estimation activities. There are many steps beyond the seed activity proposed here. Developments in this area have potential as long-term avenues for collaboration.

In the wider realm of Earth science activities at JPL, the methodology we propose to pioneer is applicable to and has potential benefit for most large-scale parallel models that are used to develop and support ocean, land, atmosphere, cryosphere and solid Earth science activities. Although these areas are beyond the scope and budget of this proposal, the development of methodology, undertaken here, can directly impact the mission life-cycle integration of state-of-the-art parallel models in all these fields.

7b. Quantitative Objectives

The ECCO2 state estimates are large computations taking several weeks to execute on several hundred to thousands of processors. The solutions that are produced contain turbulent eddies and so there is a high-degree of realistic temporal and spatial variation (see Figure 1). The realism of the solutions makes the ECCO2 solutions potentially very valuable for evaluating instrument concepts and assessing instrument design trade offs. There are innumerable interesting quantities that can be diagnosed from these solutions, but because the flow is highly variable it is not practical (or even feasible) to save all the possible combinations of terms that an end user might want to examine. For example an end user might be interested in a time series of the transport of heat across a particular density surface as a function of time in a particular region. Unless this density surface is defined in advance, it is not, in general, possible to derive this flux precisely, except by rerunning the simulation, and it is often not possible in advance to know which density surfaces will be of interest. In the context of a specific mission, for example SWOT, an end-user might ultimately want to examine correlations between this flux and observed sea-surface features. Any attempt to save enough information for all unknown analysis, therefore, falls short due to a combinatorial explosion in the amount of possible information to be stored. At present most research activities, therefore, make assumptions about the steadiness of the fluid fields, allowing approximate calculations to be made from time-averaged outputs. However, as flow becomes turbulent and inhomogeneous (i.e., as resolution increases) these assumptions become increasingly problematic.

The scale of the problem can be illustrated using Figure 1. The figure shows an instantaneous snapshot of near-surface current speeds from an ECCO2 ocean simulation with horizontal grid spacing ranging from 2 to 6.25 km. At these scales, the model explicitly computes fluxes that result from the complex interplay of mechanical and thermodynamic forcing with internal fluid dynamics. Analyzing these dynamics and mapping the behaviors to observable signals, in a realistic domain, requires analyzing correlations between multivariate data streams and analyzing combinatorial terms that characterize the turbulent interplay between fluid flow, density and pressure surfaces, heat, and salinity, as well as between wave motions and non-wave motions. It is not always possible to save sufficient information for every possible analysis and rerunning, in their entirety, computations of the sort depicted in Figure 1 can take many months and requires significant computational resources. Therefore a new strategy is proposed here.

The target for this proposal is the development of a capability that will allow exact "replaying" of sections of the ECCO2 state estimates, in order to generate time series of almost any quantity imaginable, not just those archived as part of the original state estimate. This capability will not require the computing resources of the full estimation system and will not require unrealistically large archival of every imaginable diagnostic, cross-product, and triple-product term. The proposed capability will be front-ended with a menu driven configuration engine, based on technology developed at MIT as part of separate NSF-funded activities. The result will be a demonstration system that allows extraction of arbitrary time-series of terms from the ECCO2 state estimates at arbitrary temporal frequency without requiring expert modeler involvement or overwhelming computer resources.

This will be a significant improvement on current capabilities, which are primarily based on static archives recorded from a state estimation experiment and which offer only limited possibilities for making custom analyses of high frequency phenomena. With the current approaches, long time series with high temporal frequencies are only practical for a few two-dimensional (horizontal) fields; three-dimensional time series are only available at very coarse temporal frequency, typically averaged over a month or more. Our goal is to develop capabilities to make any quantity or products of quantities available at arbitrarily high temporal resolution (up to the limits of what makes sense with the underlying numerical model, currently around one hour). Longer term, as state estimation capabilities evolve toward higher and higher resolution, our approach can scale to allow state estimates made on next-generation computer systems with thousands and even tens or hundred of thousands of processors to be queried for unexpected terms. In such systems, the possibility of saving adequate high-temporal frequency information is likely to become even less realistic and, therefore, the techniques and capabilities we propose to pioneer here will become increasingly central.

8. Approach

Planetary scale state estimates are generated by fitting a three-dimensional time-stepped simulation to observed satellite and in-situ data. State of the art high-resolution estimates, for example, those generated in the ECCO2 project, use several hundred processors and generate a decade-long time series of a standard set of terms that is approximately 10TB in volume. For the work proposed herein, we will base our development on the ECCO2 system but the technology that we plan to develop will be applicable to other, existing and future, Earth system modeling projects. The resolution of the ECCO2 calculations varies with location on the mesh, but on average it is approximately 18km. In this work we propose to augment the ECCO2 estimation output with a high temporal resolution record of instantaneous simulation state on a regular, approximately 10 degree x 10 degree mesh. This information, together with stored boundary fluxes and initial conditions will be archived as part of the state estimation process.

The highly sampled, 10 x 10 degree lines, will require a relatively modest increase in overall storage (they represent less than 3% of the total calculation grid cells). Nevertheless, once they are available they, together with a limited area model configuration equivalent to the ECCO2 model setup, provide the capability to literally "replay" an arbitrary geographic area of the calculation, sampling any calculated quantity or even some new function of model calculated quantities at arbitrarily high temporal and spatial resolution. The resulting, custom, high-frequency diagnostic terms could then be employed for subsequent input into an analysis or Observing System Simulation Experiment (OSSE)-style package. Because the results are derived from observationally constrained simulations, they contain the sort of realistic character that is important for effective observing system evaluation. Because the computational footprint of these calculations is reduced they do not require national supercomputing facilities dedicated for long periods of time.

To make the replay capabilities broadly accessible, we will adapt and deploy an XML-based model and data description approach to provide an intuitive tool for configuring and executing the limited area models. This tool can be configured to run on any Java capable platform and is currently being demonstrated in coupled MITgcm configurations (MITgcm is the ocean model underlying the ECCO-GODAE and ECCO2 assimilation activities). As part of this project we plan to demonstrate at JPL the use of this system to generate custom ECCO2 diagnostics at high frequency. The front-end system that we will employ also allows the introduction of extra code modules into the ECCO2 system. This will allow us to demonstrate the generation of unanticipated diagnostics that are a function of system state variables.

The technology we will use has been developed at MIT as part of the NSF Curator project and is designed to support non-expert users interacting with complex Earth system models. This is an ideal technology to apply to this problem. Technology from the MITgcm modeling and ECCO state-estimation activities will be employed in this work. These activities connect to numerous research efforts within the MIT Department of Earth, Atmospheric, and Planetary Sciences and to existing collaborations with JPL. MIT PI Hill will provide time on this project as part of separately funded state-estimation activities.

9. Describe the innovative features of this proposal

As far as we are aware, no capability of the sort described here currently exists as part of any Earth System modeling activity. Nevertheless, this is a potentially enormously valuable tool for extracting more value from existing Earth System information gathered by NASA, for using that information to help develop and refine next-generation observing strategies, and for validating instrument concepts. This work will hopefully lead to broader initiatives aimed at tightening integration and feedbacks between state estimates being produced today and missions being developed for the future. In this prototyping project, the connection between the replay of a state-estimate calculation and an OSSE-style workflow would not be interactive. However, once the capabilities we have developed are available it would be a natural step to consider incorporating them into an OSSE system more directly, impacting core aspects of mission evaluation.

To illustrate the potential role of the proposed technology for JPL and NASA missions and strategic goals, we consider the example of the SWOT wide swath altimeter, which is a decadal-survey-recommended mission. For the SWOT mission, there are numerous interesting questions regarding the relationship between high-resolution sea-surface height variations (that the instrument measures) and the ocean interior state. For example, a SWOT researcher might want to examine whether signals that are measured by SWOT can be used to infer sub-surface transports of dense water through critical regions such as the Denmark Strait or to monitor critical climate processes such as Deep Water formation. Using the technology that we propose to develop, it will become possible to carry out, on local computer facilities, ensemble studies using an accurate regional replay at very high-resolution of decadal-length, global-scale ocean state estimates; this regional replay can in turn be sampled and analyzed for any desired quantity. The proposed technology would make this type of analysis possible at manageable computational cost and in an interactive setting suitable for statistical engineering studies. Significantly, although our application demonstrations will be based on ocean research, the core technologies that we propose to develop is not limited to ocean science applications; the methodology can be applied to many other Earth Science OSSE scenarios, where high-end phenomena truth models are required.

In the context of the broader science community, there are many science questions, revolving around the role and accurate representation of small-scale, high-frequency processes, that the approach we propose to pioneer would enable. For example a scientist interested in studying how high-frequency, small-scale variability impacts regional physical and ecological ocean behaviors would be able to use the proposed technology to provide an accurate and realistic environment for such studies without requiring super-computer resources.

10. Contribution of Partner University

The MIT PI involved in this proposal is responsible for developing and maintaining many aspects of the MITgcm model that lies at the heart of the NASA ECCO2 and ECCO-GODAE collaborations. The MIT group is heavily involved in numerous idealized and realistic planetary scale simulation and assimilation activities. Postdoctoral researcher Evangelinos has developed the LCML, Legend tools that will provide the accessible front-end to the proposed system. Hill and Menemenlis work with colleagues at MIT, Harvard, Princeton, NASA Goddard, and NASA Ames on national programs to monitor and measure the ocean state. They are experienced in both model development and in the details of model application to a variety of problems. PI Hill has been working on developing a generic replay toolkit of the sort that will be needed for this project. Hill and Evangelinos will incorporate this toolkit into the ECCO2 system as part of this project and work with JPL partners to make the resulting system available for JPL researchers to exploit. The XML tools that Evangelinos has been developing will be customized for this project to allow JPL researchers to readily control and replay ocean state simulations with additional custom analysis.

11. Exchange of personnel

No long-term exchange of personnel is foreseen. However, Hill and Evangelinos plan two visits to JPL and Menemenlis plans one visit to MIT during this project. These visits will overlap with separately funded travel stemming from existing collaborations involving Menemenlis, Hill, and Evangelinos. During the two visits to JPL, Hill and Evangelinos will demonstrate prototype and mature versions of the proposed technology.

12. Significance and impact of results on JPL missions and programs

This project would demonstrate technology for incorporating, realistic, high-temporal and high-spatial resolution observationally constrained, Earth system models into OSSE type systems. It would also provide a world-leading technology to a broad spectrum of high-resolution ocean state estimation research and engineering work. The work would make previously inaccessible information in high-resolution state estimates more readily available. This would benefit other activities, such as regional ocean modeling process studies, which could generate custom high-temporal and high-spatial resolution ocean information for use as boundary conditions. Dynamical and biogeochemical studies of the ocean system would also benefit from the "replay" capabilities that would be developed under this proposal.

Specifically for JPL, the proposed work will allow a fuller utilization of JPL-produced ECCO2 ocean state estimation results from the science community. The proposed work has applications in the utilization and analysis of data from existing and planned oceanographic missions and programs, for example, OSTM, QuikSCAT, Aquarius, GRACE, and GHRSST. Finally, the proposed work has applications in the design of decadal-survey-recommended satellite missions, for example, DESDynI, SWOT, and GRACE-II. In particular, high-resolution results like that depicted in Figure 1 are already being provided to the SWOT science team for preliminary design purposes. Right now only the high-frequency sea surface height fields are available for the preliminary studies. The proposed work will make possible to search the full-depth, high-frequency model solution for any correlations between the surface-observed signal with sub-surface processes, as was discussed in Section 9.

13. Has the proposal been submitted elsewhere?

No

14. Plans for follow-on funding

Follow-on funding for this work will be requested from future NASA calls and from equivalent NSF programs. PI Hill is also working with colleagues at an industry partner (VMWare) to develop virtual machine technology that would provide ongoing near-seamless support for technology approaches of the sort proposed here. We anticipate proposing to build on this work through a follow on proposal to the NASA AIST program, which supports advanced, multi-disciplinary, information system technology work of this nature for the science mission directorate. We also expect to approach the NSF STCI program, which funds strategic technology development for cyberinfrastructure.

15. Budget
See budget sheet below.

16. Partner contract administrator contact information
Amy Alajajian, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Building E19-750, Cambridge, MA 02139-4307; phone: 617-253-2495; email: afavalor@mit.edu (see attached letter).

17. JPL Principal Investigator Signature

Name: Dimitris Menemenlis — Org: 3248

Signature: D. Menemenlis — Date: 12-11-2008

17a. JPL PI Section Manager Signature

Name: Yi Chao — Org: 324

Signature: — Date: 12-11-08

18. JPL PI Division Manager Signature

CR

Name: Harold Yorke — Org: 32

Signature: Harold W. Yorke — Date: 12/11/08

19. University Co-Investigator Signature

Name: Christopher Hill

Signature: Chris Hill — Date: 12-5-08

20. University Representative with Signature Authority

Name: Michael P. Corcoran

Title: Assistant Director, Office of Sponsored Programs

Signature: — Date: 12/8/08

Budget Sheet

Category	AT JPL	AT EXTERNAL INSTITUTION(S)
DIRECT COST		
1. Salaries (Itemize) *(Only "itemize" the person names or job classifications and the number of hours for each. You can show one total $ salary figure for labor.)*	D. Menemenlis, 75 hrs $4,300.00	C. Evangelinos, 1160 hrs $36,087.00
2. Labor Fringe Rates - Employee Benefits	$2,100.00	
3. Cat A Labor (Itemize) *(Only "itemize" the person names or job classifications and the number of hours for each. You can show one total $ figure for labor.)*		
4. Procurements –Equipment, Materials and Supplies (Itemize). JPL - Do not list the contracts for outside collaborators. This total is on line #12 on the external collaborator column.		
5. Procurements – Subcontracts (Itemize) *(PS – contracts other than with collaborators)*		
6. Services – (Itemize) *(JPL be sure to include in-house services at JPL)*		
7. Domestic Travel *(only as a research cost; and domestic conference travel is allowed up to a maximum 5% of the total budget*) Itemize with what and where the travel is required.	1-week domestic travel to MIT/Cambridge $2,400.00	
8. Other (Itemize) *(Chargebacks, etc.)*	$300.00	
9. Total Direct Costs (total of dollars 1 through 8)	$9,100.00	$36,087.00
10. ALLOCATED DIRECT COSTS (ADC)	$11,700.00	
ADC **FY09** - See Section Administrator or Business Administration Manager for current rates. ADC costs are calculated on the JPL's total direct costs Item #9 and the external institution(s) budget item #12.		
ADC at JPL consisting of: a. Labor ADC b. RSA Contract ADC c. Other Contracts ADC d. Purchase Orders e. General ADC Enter total on Item #10	$2,100.00 $8,300.00 $1,300.00	
11. Overhead -external Institution		$32,023.00
12. Individual Budget: *(JPL add Item #9 Direct Cost and #10 ADC costs for total JPL budget)* *External Institution add Item #9 and Item #11 Overhead for total)*	$20,800.00	$79,117.00
13. Combined Budget: *(JPL Budget plus External Institution Budget)*	$99,917.00	

Figures, Graphics, Tables, etc.

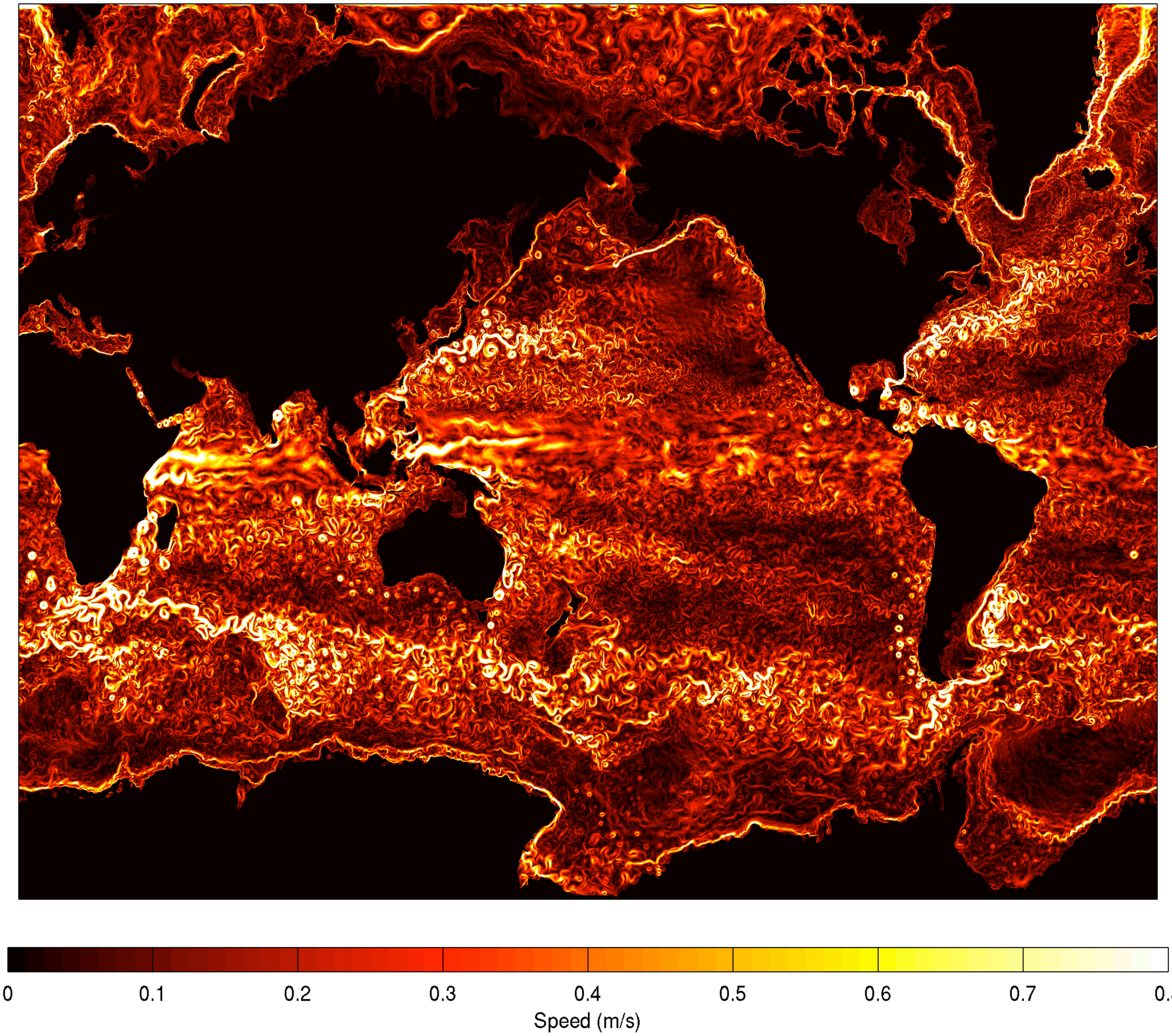

Figure 1: Snapshot of near-surface (15-m depth) current speed on March 3, 1993 from a quasi-global simulation with 1/16th-degree horizontal grid spacing (horizontal grid spacing is approximately 6.9 km at the Equator and 1.19 km at +/-80°). High-resolution simulations are being used by the ECCO2 project to estimate model errors and to inform sub-grid-scale parameterizations (Hill et al., 2007).

Amy E. Alajajian
Grant & Contract Administrator

Massachusetts Institute of Technology
77 Massachusetts Avenue, Building E19-750
Cambridge, Massachusetts 02139-4307

Office of Sponsored Programs

Phone 617.253.2495
Fax 617.253.4734
Email afavalor@mit.edu
http://web.mit.edu/osp/www/

December 9, 2008

Yi Chao
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91109
Subject: NASA Subcontract Proposal

Dear Yi Chao:

Massachusetts Institute of Technology formally submits herewith a proposal entitled "Strategic Technology Research to Improve the Utilization of Assimilated Ocean State Estimates in Future Mission Design and Planning and to Seed New Collaborations in this Area" The proposed research would be performed by MIT's Department of Earth, Atmospheric & Planetary Sciences. Christopher Hill would serve as Principal Investigator.

Our estimate of the cost for the period January 1, 2009 through December 31, 2009 is $79,116.00.

Should this proposal be chosen for award by the National Aeronautics and Space Administration (NASA), MIT is prepared to negotiate a cost reimbursement agreement with JPL, terms and conditions appropriate for a non-profit educational institution. Please understand that any acceptance of a resulting award for this program is contingent upon acceptable negotiations. MIT's policies prohibit the acceptance of awards that include provisions that restricts our ability to disseminate research results or place limitations on the use of foreign nationals working on research. MIT must be able to publish research results without restriction and must retain title to all intellectual property (with a non-exclusive license to the government).

Questions relating to technical aspects of this proposal should be directed to the principal investigator of this proposed effort. Questions of an administrative nature should be addressed to the undersigned.

Sincerely yours,

Amy

Amy Alajajian
Grant & Contract Administrator